Pre-Analysis Plan: Improving the workplace experience of apprentices

For details of the trial design and power calculations, see the pre-registration for this trial.

Outcome measures

Primary outcome measure

The primary outcome measure is apprentice retention, as a binary measure (whether the apprentice is still in a training contract during trial). We will collect data on apprentice retention at four months after the first intervention message is delivered.

As it could take several months for retention data to be updated in the apprentice database, retention data at four months will be re-extracted seven months after delivery of the first message of the last batch to obtain more accurate retention data. See the 'Trial Design' section for further details for the timing the primary outcome measure.

Likelihood of contract cancellation (median time to drop-out)

This will use retention data and will involve time to drop-out survival analysis, using contract drop out as 'failure'. We will look at median time to drop out at the end of follow-up period between the control and the treatment group and will also estimate and compare survival functions between the treatment and the control group.

Other quantitative outcome measures

Data for other quantitative outcome measures (below) will be collected from surveys of apprentices and their employers in both the treatment and the control groups. The surveys will be conducted at the end of the intervention period.

Apprentices' self-reported supervisor support

This is a 4-item survey measure with 5 response categories (Never, Rarely, Sometimes, Often, Always. The scores range from 1 (Never) to 5 (Always). The final supervisor support score will be the mean score of the 4 items.

Apprentices' self-reported workplace experience

This is a 7-item survey measure. The scores range from 1 (Never) to 5 (Always) with higher scores indicating a better workplace experience. The final workplace experience score will be the mean score for these seven items.

Apprentices' and employers' awareness and understanding of workplace laws

Awareness is measured using a single item in the survey. There are other indicators of awareness included in the apprentice survey (e.g. whether they know which award they are covered under and whether they know if they are being paid correctly etc.).

Apprentices' and employers' confidence to raise and address workplace issues

Two-item measure in the surveys measure participants' confidence to raise and address workplace issues.

We will also undertake subgroup analyses (see 'Randomisation' and 'Analysis' sections).

Descriptive and qualitative outcome measures

We will also be examining perceived social norms on pay and conditions (7-item construct), employer reported workplace experience of their apprentices, differences in perceptions of workplace experience between employers and apprentices. However, we do not intend to undertake hypothesis testing but instead undertake descriptive and qualitative analyses (see 'Qualitative analysis' section).

Apprentice workplace experience and the mechanism of impact of intervention messages will be examined qualitatively through survey and interview data.

Data sources

This trial will use and collect data from a variety of sources and they are summarised in Table 1.

Retention data will be provided by the Department of Employment, Skills, Small and Family Business (Employment) who routinely collects and holds data on Australian apprenticeships.

We will conduct an online survey of apprentices and employers to obtain data on their workplace experience. We will also conduct ten interviews with apprentices in the treatment group (five who remained in contract, five who left their contract) to obtain a deeper understanding of their experience and feedback on the intervention messages. Ten employers in the treatment group will also be interviewed.

Table 1. Data for the trial

Outcome measure(s)	Type of data	Data source
Retention (primary outcome)	Administrative data (apprentices training contract data)	Employment
Supervisor support	Survey data (survey of apprentices)	BETA/FWO survey
Workplace experience of apprentices	Survey data (survey of apprentices)	BETA/FWO survey
Awareness and understanding of workplace laws	Survey data (surveys of apprentices and employers)	BETA/FWO surveys
Perceptions of social norms around pay and conditions	Survey data (survey of employers)	BETA/FWO survey
Confidence in dealing with workplace issues	Survey data (surveys of apprentices and employers)	BETA/FWO survey
Likelihood of contract cancellation (median time to drop out)	Administrative data (apprentices training contract data)	Employment
Feedback on intervention messages and action taken, workplace experience	Semi-structured interviews	BETA interviews

Hypotheses

Primary hypotheses

We hypothesise that the apprentice retention rate in the treatment group will be higher than the rate in the control group.

We hypothesise that the risk of contract cancellation at the end of the follow-up period will be lower for apprentices in the treatment group compared to apprentices in the control group.

Other hypotheses

We further hypothesise that:

- apprentices in the treatment group will report better supervisor support compared to apprentices in the control group
- apprentices in the treatment group will have a better workplace experience compared to those in the control group.
- the risk of training contract cancellation at the end of the follow-up period is lower in the treatment group compared to the control group
- employers in the treatment group report a higher level of awareness of workplace laws compared to employers in the control group.
- apprentices and employers in the treatment group report a higher level of confidence in raising and resolving workplace issues compared to those in the control group.

All our hypotheses are directional because we are interested in whether the intervention leads to improved retention, awareness and understanding, confidence in raising workplace issues and improved workplace experience for apprentices.

We also have the following research questions which will be answered using descriptive and qualitative data.

- Do employers in the treatment group report better workplace experience for their apprentices compared to employers in the control group?
- Are there differences in self-reported social norms around pay and conditions among employers in the treatment and the control group?
- Do perceptions of workplace experience differ between apprentices and employers overall?
- Do perceptions of workplace experience differ between apprentices and employers in the treatment group and the control group?

We are sending intervention messages via SMS for apprentices and emails to employers (preferred channels). Alternative channels may be used if the contact details are missing or incorrect. This means messages will be sent to participants via one channel only. In this trial, we are not making any hypotheses about communication channel nor testing the effect of the channel of communication as the trial is not designed and powered to detect the effect of communication channels.

Randomisation and Trial Procedure

This is a two-arm cluster randomised controlled trial with clustering at the business level. Randomisation will be stratified by business size. We will conduct balance checks after randomisation.

Apprentices and their employers will be enrolled in batches, based on monthly commencement data of apprentices, to reach the desired sample size.

Trial threats

Blinding

Apprentices and their supervisors/employers in the treatment group will be aware that they are receiving educational messages from FWO but they will not be told that they are part of a trial. Therefore, we believe there is no threat of bias or contamination in this scenario.

Spillovers

We think the chance of spillovers is low as we are clustering at the business level to prevent spillovers between apprentices in the same business. However, it is possible that apprentices in the treatment group could speak to apprentices in the control group at their off-the job training (e.g. TAFE) and this could influence the retention rate of apprentices in the control group. We think this is unlikely but will measure this as part of our survey of apprentices. While this will not be a perfect measure, it will give us some indication of the extent of any spillovers.

Attrition/missing data

For retention outcome, we expect little, if any missing data as we are using administrative data. Participants can opt out at this stage from the trial but expect opt out rates from administrative

data to be low. If the participant has opted-out (identified via privacy notice opt-out flag) before randomisation, we will exclude them from the trial.

For survey data, it is likely that the response rate will be higher in the treatment group than the control group. While we will optimize our survey design to improve response rates and use reminders to encourage completions. Those who opt-out of intervention messages will not be further contacted for surveys or interviews so they will not be included in survey/interview data and analysis.

We will test for differential attrition and whether data is missing completely at random (MCAR), missing at random (MAR) or not missing at random (NMAR). Missing data mechanism may differ between the treatment and control group for survey data. If the data is MAR, we will re-weight the data with the probability of responding to the survey (Lohr, 1999; Little & Rubin, 2002). If the data is NMAR, we will use the heckman selection model (Heckman, 1976; (Heckman 1976, Allison 2001). We will also construct Manski-Horowitz bounds around estimates as we cannot be fully sure about the underlying missing data mechanism for survey data and whether any missing data approach will fully address bias (Manski 1990, Horowitz and Manski 1998, Horowitz and Manski 2000).

Analysis

Main analysis

For all outcome measures, we will undertake intention to treat (ITT) analysis as our primary analysis.

Baseline covariates that will be included in main analyses include apprentice age group, apprentice gender, and business type and month as a dummy variable. However, if these covariates and strata variables are not improving the precision of our estimates, we will not include them for our main results but instead report results with covariates and strata in an appendix for comparison or in a supplementary technical report (to be determined at a later stage).

Retention rate

For the retention rate outcome measure (Y), we will adjust for clustering at the employer level and strata and baseline covariates. We will undertake linear regression analysis:

$$Y_{ij} = \alpha + \beta T_i + \gamma X_i + v_j + \omega_{ij}$$

T indicates whether it is the treatment or the control group. X_i indicates the set of mean-centred covariates, including strata that could be included in the linear model. These covariates will also be interacted with the treatment variable. v explains the group level error term and ω is the individual level error term. We will calculate robust standard errors and confidence intervals. We will also conduct logistic regression as a robustness check for this and other binary variables in our trial. We will also undertake a random effects logistics model as a robustness check of retention rate primary outcome measure.

We will undertake survival analysis and compare the median time taken to drop out of apprenticeship between the treatment and the control group.

We will also undertake ITT analyses for survey outcome measures (to test supplementary hypotheses) if we achieve response rates of 25% or more. For apprentice level outcome measures, we will adjust for clustering at the employer level as well as strata. For employer level outcomes, there will be no adjustment for clustering. Analyses will be conducted with those who completed the survey only. We will investigate differences between responders and non-responders as this will have implication for the generalizability of survey findings.

We will perform a one-sided test as indicated under Power Calculations.

Subgroup analyses

We will conduct subgroup analyses by apprentice's gender, apprentice's age group, history of apprentice retention (based on contract cancellations in past 5 years but this data will only be reliable for established businesses) and business type, without adjusting for covariates to improve sample size for each regression model. Separate models will be conducted for the sub-groupings.

Adjustments for multiple testing will not be undertaken for several reasons. First, our primary outcome measures are clearly stated, defined and the two primary hypotheses are complementary. Second, the main theory we are testing with the primary hypotheses is that our intervention leads to an improvement in retention. If the results of our two primary hypotheses are contradictory, we will interpret our treatment as having an effect on retention if retention rates significantly differ between the treatment and the control group (primary hypothesis 1). Third, as mentioned previously, we will not undertake sub group analyses if the primary hypotheses are not supported, thereby reducing the chance for false positives. Lastly, we are not viewing alpha of 0.05 as the strict threshold but will be using confidence intervals and other statistical and contextual information (see 'Interpretation of results' section for more information). For more information on the discussion on multiple testing and the use of p-values and the new approach to statistics, please see papers from Rothman (1990), Althouse (2016), Cumming (2014), Wasserstein and Lazar (2016).

Supplementary analyses

Apprentice's working conditions will be examined as part of secondary analyses. Outcome variables will be treated as continuous variables. Adjustment for clustering, strata and covariates will be undertake as per the main analyses.

While our main outcome measure is contract retention rate with a particular employer, from vocational educational policy perspective, an apprentice moving to a different employer but continuing their apprenticeship is still a positive outcome. This will be examined as part of our secondary analyses as the accuracy of this data will vary depending on how long an apprentice has been followed during the trial.

We will also calculate complier average causal effect (CACE) for our primary outcome measure. This is because we expect that not everyone in the treatment group will receive our intervention as intended. If contact details in the apprentice administrative data are incorrect or if the employer or apprentice has changed their contact details following training contract registration, we will not be able to deliver intervention messages to these participants. This will be estimated using two-stage least squares regression. For this analysis, if an apprentice or employer receives at least one intervention message, they will be considered as having received the treatment. This will include anyone who has opted-out after the first intervention message as well as those who don't opt-out but have not opened the message. This is a conservative approach and will result in an underestimate of the CACE.

As exploratory analysis, we will also conduct dose-response analyses for the primary outcome measure. This will provide information on the relationship between the number of messages received, the type of recipient and retention. It will also allow us to examine the shape of the dose response curve. In short, it will tell us whether more messages lead to greater retention or whether a single message is sufficient. It can also tell us whether it is the pairing of the messages that makes it more effective or whether sending messages to an individual type of recipient (e.g. employers only) is as effective. As there is likely to be selection bias as participants can opt-out of receiving messages (McGowan, Nix et al. 2010), we will estimate dose-response effects using instrumental variable analysis as per Maracy and Dunn (Maracy and Dunn 2011).

Table 2. Dose-response analysis

	Analysis 1a	Analysis 1b	Analysis 1c
Description	Dose response analysis by recipient – main effect Employer/supervisor	Dose response analysis by recipient – main effect apprentice	Dose response analysis by recipient – interaction effect of 1a and 1b
Lowest dose	Dose of 1 (the minimum dose)	Dose of 1 (the minimum dose)	
	2	2	
	3	3	
	4 (maximum dose)	4 (maximum dose)	

Interpretation of results

We will make use of p-values to aid in the interpretation of our results. However, we will avoid taking a 'bright line' approach, in which a threshold is used to determine a meaningful finding. Instead, we will adopt a 'fuzzy line' approach and will consider the p-value together with prior evidence, effect size, outcome variability and design limitations in order to assess the strength of a finding.

Our intervention of delivering text and email messages is a relatively low cost intervention and once trialled, we think it will be simple to scale up. The risk of adverse effect from our information intervention is low. Given these reasons, we are less worried about having a false positive finding. If we detect effect sizes that are material and are of practical significance to FWO, we will still take effects seriously even when p-values are above the specified threshold.

Qualitative evaluation

Interviews will be audio-recorded if permitted, transcribed and analysed in NVivo. Thematic analysis will be undertaken. Findings from qualitative interviews will be supplemented with findings from survey data to provide additional context for findings.

Pre-analysis plan commitments

No analysis has been undertaken prior to the completion of this pre-analysis plan.

We will be transparent about, and provide justification for, any deviations (additions or omissions) from this plan.

References

Allison, P. D. (2001). Missing data, Sage Publications.

Althouse, A. D. (2016). "Adjust for Multiple Comparisons? It's Not That Simple." The Annals of Thoracic Surgery **101**(5): 1644-1645.

Cumming, G. (2014). "The new statistics: Why and how." Psychological Science **25**(1): 7-29.

Heckman, J. J. (1976). The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. Annals of Economic and Social Measurement, Volume 5, number 4, NBER**:** 475-492.

Horowitz, J. L. and C. F. Manski (1998). "Censoring of outcomes and regressors due to survey nonresponse: Identification and estimation using weights and imputations." Journal of Econometrics **84**(1): 37-58.

Horowitz, J. L. and C. F. Manski (2000). "Nonparametric analysis of randomized experiments with missing covariate and outcome data." Journal of the American statistical Association **95**(449): 77-84.

Manski, C. F. (1990). "Nonparametric bounds on treatment effects." The American Economic Review **80**(2): 319-323.

Maracy, M. and G. Dunn (2011). "Estimating dose-response effects in psychological treatment trials: the role of instrumental variables." Statistical Methods in Medical Research **20**(3): 191-215.

McGowan, H. M., et al. (2010). "Investigating the impact of selection bias in dose-response analyses of preventive interventions." Prevention Science **11**(3): 239-251.

Rothman, K. J. (1990). "No Adjustments Are Needed for Multiple Comparisons." Epidemiology **1**(1): 43-46.

Wasserstein, R. L. and N. A. Lazar (2016). "The ASA's Statement on p-Values: Context, Process, and Purpose." The American Statistician **70**(2): 129-133.